alteryx

WHITEPAPER

Private Data Handling

CONTENTS

PRIVATE DATA HANDLING 3

PRINCIPLES 3

PURPOSE 3

PRIVATE DATA HANDLING 4

- Prerequisites 4

DATA SECURITY 5

- Architecture 5
- Example Flow 6
- Data Handling 7
  - Inventory and Classification 7
  - Data: Return, Destruction & Disposal 7
  - Encryption in Transit 8
  - Encryption at Rest 8
  - Metadata Collection 8

DEPLOYMENT 9

- Separation of Duties 9
- Patching 9
- Vulnerability Management 9
- Third-Party Components 9

IDENTITY AND ACCESS MANAGEMENT 10

- Requirements 10

CREDENTIAL MANAGEMENT 10

AUDIT 10

Private Data Handling

The following paper provides a preview of the security controls, architecture, and governance practices utilized to secure an Alteryx Analytics Cloud Platform private data plane. This is a data processing cluster that customers can run on their own infrastructure to keep data workloads within their own networks.

Principles

Alteryx powers analytics for all by providing the leading analytics cloud platform. Alteryx delivers easy end-to-end automation of data engineering, analytics, reporting, machine learning, and data science processes, enabling enterprises everywhere to democratize data analytics across their organizations for a broad range of use cases. Our Alteryx Analytics Cloud Platform is supported by a robust and comprehensive Information Security Program, which is based on a defense-in-depth philosophy and a secure by default approach. This program is designed to protect our platform and our customers' data against potential threats, and to ensure the integrity and reliability of our solutions. By implementing these security measures, we provide our customers with the confidence and peace of mind to trust and rely upon our platform to support their data analytics needs and accelerate digital transformation.

By employing experts in information, cyber, network, and application security, Alteryx demonstrates its commitment to continually securing the company’s defensive systems, ensuring consistent security review processes, maintaining our secure infrastructure, and assuring adherence to our Security Policies. Security is designed into our products and operations.

We believe when products are built in secure environments, we create a more secure world to work out of, and our customers directly benefit from these protections. Focusing on security and data protection are among our primary design criteria.

Purpose

At Alteryx, we are committed to maintaining the highest standards of security and transparency. We are dedicated to staying ahead of the constantly evolving security landscape, and to implementing the latest industry-specific standards and best practices to protect our customers' data. Our comprehensive security measures are designed to prevent accidental or unlawful destruction, loss, alteration, unauthorized disclosure, or unauthorized access to data.

This whitepaper is intended to provide an overview of our security practices in the Alteryx Analytics Cloud Platform when configured for private data handling and to demonstrate our commitment to ensuring the integrity and reliability of our products, environment, and data. For more information about our overall security and privacy programs, please visit alteryx.com/trust. We look forward to working with you and to helping you achieve your business objectives while protecting your data.

Private Data Handling

Private Data Handling refers to the overall architecture of the Alteryx Analytics Cloud Platform that provides more flexibility for data handling and storage. In simple terms, the portions of the Alteryx Analytics Cloud Platform that handle the customer's internal data are deployed in a manner that gives the customer complete control over where their data is stored and processed. We utilize industry best practices to build comprehensive defense in depth security controls to give you confidence and meet compliance obligations.

This is accomplished by creating multiple planes, which we call the control plane, or the portion of the architecture that resides in Alteryx's infrastructure, and the data plane, which resides in your cloud infrastructure. The control plane is responsible for user interface interactions, deployment capabilities, monitoring, storage of application metadata, and communication brokering. The data plane is responsible for establishing connections to data sources, the processing of data, and storage of uploaded datasets and data samples at rest.

Third-party data storage or SQL execution engines such as Snowflake, Databricks, or BigQuery are classified as part of the data plane (where customer content is stored and processed), but those are not the focus of this paper.

Prerequisites

To support the deployment of the data plane infrastructure, a cloud account must be created with a VPC (Virtual Private Cloud) in a region of your choosing. Connectivity must also be created between the data plane infrastructure and any other environments necessary to support any workload use cases. Data plane infrastructure is deployable to an Amazon Web Services (AWS) environment.

Data Security

Alteryx has taken great care to provide for the security and integrity of data, regardless of where it resides or is transmitted within the Alteryx Analytics Cloud Platform infrastructure. Alteryx classifies data into two main categories: customer content and application metadata. Customer content belongs to the customer and refers to the data that is being blended, transformed, analyzed, or processed. Application metadata is the data that Alteryx uses to correctly process customer content as directed.

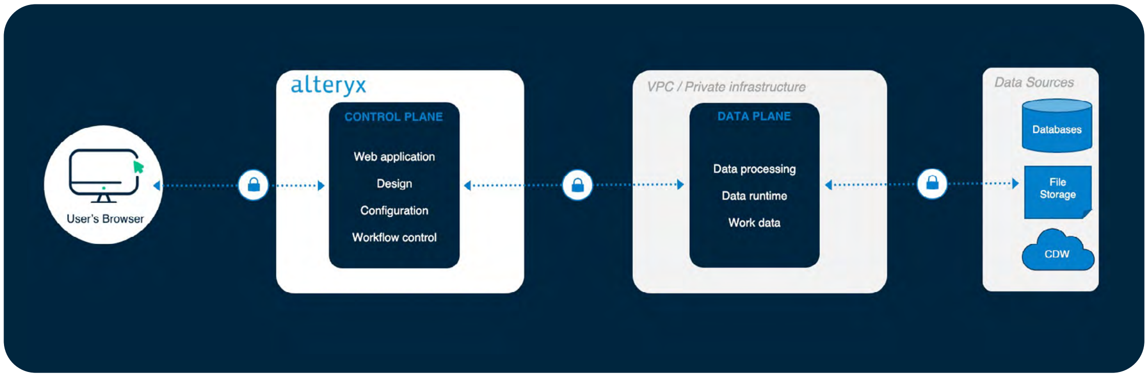
*Throughout this paper, references to “data” will refer to customer content, unless otherwise noted.

Architecture

The data plane segregation from the Alteryx environment as well as other customer environments prevents co-mingling of data and helps ensure that no cross-tenant data leakage can take place.

This network design approach separates the control plane, which is responsible for routing and forwarding decisions, from the data plane, which is responsible for forwarding data packets. This separation helps to improve network security by isolating the control plane from external attacks and preventing malicious actors from gaining access to sensitive routing information. Additionally, this type of architecture allows for the use of different security measures and controls on the control and data planes, which can further enhance security posture.

The communication flows through the system are segregated such that control-related messages are directed to the Private Data Handling infrastructure, and return messages contain only metadata about requested actions and sample data for display purposes. Sample data is not centrally cached nor stored. Network communication over the public internet is encrypted with TLS 1.3 or higher. For the connection between the private data plane and the data source, this remains true if supported by the data source.

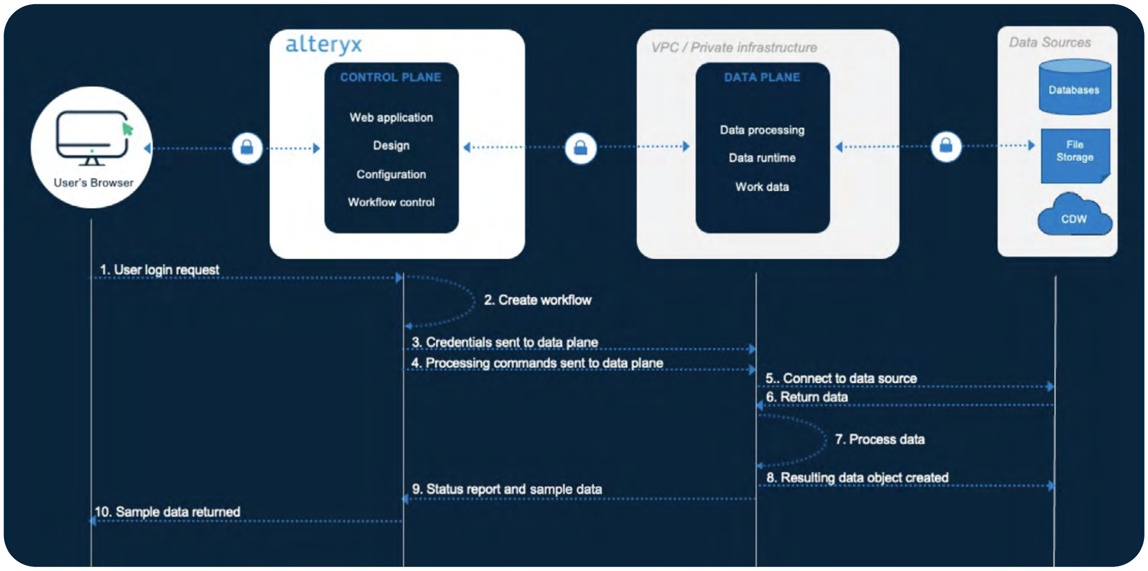


Example Flow

For example, let us assume that you are a data engineer, and you want to retrieve records from a cloud database, normalize the data into a dataset, and store the resulting records in an S3 bucket.

In this example, it is assumed that a workspace is configured with Single Sign-On (SSO) integration, a data plane has successfully been deployed to your environment, and configuration of database credentials and IAM roles necessary have already been set up.

1. The flow begins with you logging into the Alteryx portal using your identity provider for SSO authentication. Native authentication may also be used in instances where an external identity provider is not available.
2. You then utilize a configured database connection as part of your workflow and create a transformation of the resulting data to be stored in a configured S3 bucket.
3. The credentials configured for the database have envelope encryption applied and are sent to the data plane for connecting to the data source.
4. The processing commands based on the workflow design are also sent to the data plane.
5. The data plane decrypts the credentials utilizing keys stored as secrets (also using envelope encryption) and initiates the connection to query the data.
6. The data results are sent to the data plane for processing.
7. The private data plane processes the data according to the workflow.
8. The private data plane then utilizes IAM credentials retrieved with an IAM role to create an object in S3 with the resulting data.
9. The data plane reports back status and sample data to the control plane.
10. Sample data is then streamed back to your browser.



Data Handling

At Alteryx, we have implemented a set of standard policies and practices for data inventory, classification, retention, and storage that apply to our overall Analytics Cloud Platform architecture and the deployment of Private Data Handling environments. This helps to ensure that data is easily accessible and can be quickly retrieved when needed, while also protecting it against unauthorized access or tampering. Our commitment to standardization helps to provide our customers with a consistent and reliable experience when using our Analytics Cloud Platform environments and ensures that their data is properly protected and managed.

Inventory and Classification

Policies regarding the appropriate tagging of collected data according to the designated sensitivity are in place to ensure data is appropriately tracked and maintained.

Data: Return, Destruction & Disposal

At Alteryx, we have implemented a comprehensive set of policies and procedures for the disposal of tangible and intangible property, including data. These policies are designed to ensure that data is tracked, reviewed, maintained, and disposed of in a way appropriate to its sensitivity and defined data retention periods. This means that data is only retained for the period required by applicable legal and contractual requirements, after which it is securely disposed of. By following these policies, we can ensure that customer content is properly handled and protected throughout its entire lifecycle. This helps to maintain the security and integrity of the data and prevents potential breaches or unauthorized access. Our commitment to secure data disposal helps our customers with peace of mind that their data is handled responsibly and securely.

Encryption in Transit

To help ensure the security of data in transit within Alteryx systems, including the Private Data Handling infrastructure deployed in the customer environment, we utilize Transport Layer Security (TLS) version 1.3 or higher.

This encryption protocol helps to prevent unauthorized access to data as it is transmitted between different components of the system. Additionally, control plane to data plane communication and intra-cluster communication are both encrypted to further protect customer content. These measures ensure that data remains secure as it is transmitted within the Alteryx system and prevents potential vulnerabilities or breaches. By implementing encryption in transit, we can provide our customers with the peace of mind that their data is protected and secure.

Encryption at Rest

Data stored within a workspace is encrypted at rest by default utilizing 256-bit Advanced Encryption Standard (AES) block ciphers, and cloud native services for key management versioning and rotation.

- Secrets within the data plane cluster are encrypted using envelope encryption separate from workloads supported by customer-managed keys.

- Customers can use encrypted storage buckets and databases as data sources.

Customer data sources outside of the workspace will adhere to the underlying encryption policy for that data source. For example, if a customer connects to an S3 bucket as an external data source, any files there will follow the customer-defined policy.

Metadata Collection

A set of application monitoring usage data is collected from the Private Data Handling environment to monitor and maintain the system’s operational stability. This metadata is a small sampling of the overall workflow data and is carefully aggregated, obfuscated, and then utilized to provide valuable insights into the system's performance. This helps to ensure that the Private Data Handling environment is running smoothly and efficiently and allows for proactive identification and resolution of any potential issues or bottlenecks.

By regularly monitoring and analyzing the metadata, Alteryx ensures our customers’ environments are stable and reliable, while allowing us to make informed decisions about how to optimize and improve their performance.

Alteryx also collects a set of application usage data through the browser as end-users interact with the application. Like monitoring data, this usage data is also aggregated and obfuscated. Alteryx analyzes this metadata for insight into how end-users interact with the app, utilization of features, and opportunities to improve the end-user experience.

Deployment

Separation of Duties

Alteryx creates the necessary infrastructure to run data plane management within a customer's cloud account. No existing infrastructure, such as network and storage, is modified during this process. All necessary infrastructure created during deployment complements any existing infrastructure, and no changes are made to governance policies in the environment.

Patching

Clusters in the data plane are updated regularly with security-patched images. Patching of components in the data plane is accomplished using a continuous delivery flow, which ensures that updates are quickly and seamlessly deployed to the system. This process begins when updates are pushed to the Alteryx repositories, which contain images that have undergone thorough testing and verification. When a new update is available, it triggers a deployment to the private data plane, which installs the update and ensures that the data plane remains up-to-date and secure. This continuous delivery flow allows for efficient and effective data plane patching and helps prevent security vulnerabilities and maintains the system's stability and reliability. By implementing this process, Alteryx ensures that the data plane is always up-to-date and protected against potential threats.

Vulnerability Management

Workload images deployed utilize the same secure Software Development Life-Cycle (SDLC) practices we utilize across the company. Containers begin with hardened baseline images and are scanned for any potential vulnerabilities.

For more information about our SDLC practices or information security program visit alteryx.com/trust.

Scanning at Alteryx includes static, dynamic, software composition, infrastructure code, and image vulnerability analysis.

Security patching of workload images is determined by the criticality of any potential vulnerabilities that might exist, and the process follows the Alteryx vulnerability policy. In general, images are pushed using rolling updates within the cluster.

Cloud services utilized in the private data plane are patched according to the schedules set by AWS.

Third-Party Components

The following describes third-party workload images utilized for operations within the data plane. Images utilized from the vendor are always patched with the latest security updates.

COMPONENT	USAGE
ArgoCD	Management of cluster workloads
DataDog	Relays cluster metrics and telemetry
Teleport	Remote connections for support

Identity and Access Management

Requirements

The Alteryx Analytics Cloud Platform integrates with AWS IAM with a least-permissions model for services and users to get appropriate access. An IAM (Identity and Access Management) account and an associated role with a required set of permissions must be provided to Alteryx by the customer to support initial deployment.

This account and role are utilized for initial deployment purposes and for any necessary changes to the data plane infrastructure needed by Alteryx. Configuration scripts are available to automatically create the necessary permissions.

Credential Management

Connections to various systems for building datasets are initiated by the data plane only. Credentials are stored utilizing envelope encryption. A data encryption key (DEK) is used to encrypt the credentials, while a key encryption key (KEK) is used to encrypt the key. This encryption is used for both storage and while in transit.

The data plane integrates with cloud-native Identity and Access Management services to utilize roles for data access without the need for storage of credentials in the environment.

Audit

Because the data plane runs within your environment, the governance and configuration of logging you have applied will also apply to the data plane. Events that occur such as infrastructure changes and account access will be logged and reviewable according to your logging policy.

alteryx

Analytics for all employees, systems, and decisions. Accelerate your analytics journey with the Alteryx Analytics Cloud Platform.

Explore Now

The Alteryx Analytics Cloud Platform empowers you to tackle your toughest business problems with powerful and easy-to-use analytics in one platform. Quickly share workflows with ease across cloud, desktop, and on-prem. You can gain quick insights with cloud-native analytics all in the browser. Get ready to experience seamless out-of-the-box integrations with modern cloud ecosystem applications. Utilize Alteryx Analytics Cloud Platform to surface hidden signals and focus on what matters with AI-driven, automated insights, and impact outcomes with built-in recommendations and best practices shaped by the Alteryx Community.

alteryx.com